

SEQUENCE LISTING

<110> Gellissen, Gerd
Braus, Gerhard
Pries, Ralph
Krappmann, Sven
Strasser, Alexander

<120> Nucleic Acid Molecule Comprising a Nucleic Acid Coding for a Polypeptide with Chorismate Mutase Activity

<130> 029474-5007-00

<140> 10/042059

<141> 2001-10-25

<150> DE 199 19 124.7

<151> 1999-04-27

<160> 7

<170> PatentIn version 3.1

<210> 1

<211> 843

<212> DNA

<213> Hansenula polymorpha

<220>

<221> CDS

<222> (1) .. (843)

<223>

<400> 1

atg gac ttt atg aag cca gaa aca gtg ctg gac ctt ggc aac att aga	48
Met Asp Phe Met Lys Pro Glu Thr Val Leu Asp Leu Gly Asn Ile Arg	
1 5 10 15	

gat gcc ttg gtc cgg atg gag gat acg atc atc ttc aac ttt atc gag	96
Asp Ala Leu Val Arg Met Glu Asp Thr Ile Ile Phe Asn Phe Ile Glu	
20 25 30	

cgg tcg cag ttc tat gcg tcg ccc tcg gta tac aaa gtc aac cag ttc	144
Arg Ser Gln Phe Tyr Ala Ser Pro Ser Val Tyr Lys Val Asn Gln Phe	
35 40 45	

cct att ccc aac ttc gac ggc tcg ttc ttg gac tgg ctg ttg tcg cag	192
Pro Ile Pro Asn Phe Asp Gly Ser Phe Leu Asp Trp Leu Leu Ser Gln	
50 55 60	

cac gag cga atc cat tcg cag gtg agg aga tac gac gcg cca gac gag	240
His Glu Arg Ile His Ser Gln Val Arg Arg Tyr Asp Ala Pro Asp Glu	
65 70 75 80	

gtg cct ttt ttc ccc aac gtg ctg gaa aaa acg ttt ctg ccc aag atc	288
Val Pro Phe Phe Pro Asn Val Leu Glu Lys Thr Phe Leu Pro Lys Ile	
85 90 95	

RECEIVED
TECH CENTER 1600/2900
02 SEP 11 PM 1:56

aa

aac tac cca tcg gtg cta gcc tcc tac gcg gat gaa atc aac gtc aac Asn Tyr Pro Ser Val Leu Ala Ser Tyr Ala Asp Glu Ile Asn Val Asn 100 105 110	336
aaa gag ata ctc aag atc tac acg tca gag ata gta cca gga ata gct Lys Glu Ile Leu Lys Ile Tyr Thr Ser Glu Ile Val Pro Gly Ile Ala 115 120 125	384
gca ggc agc gga gag cag gag gac aac ctt ggc tcg tgc gca atg gcc Ala Gly Ser Gly Glu Gln Glu Asp Asn Leu Gly Ser Cys Ala Met Ala 130 135 140	432
gac atc gag tgc ctg cag tcg cta tcc aga aga atc cat ttt ggc cgt Asp Ile Glu Cys Leu Gln Ser Leu Ser Arg Arg Ile His Phe Gly Arg 145 150 155 160	480
ttt gtc gca gag gct aaa ttt atc agt gag ggg gac aag att gtg gat Phe Val Ala Glu Ala Lys Phe Ile Ser Glu Gly Asp Lys Ile Val Asp 165 170 175	528
ctg atc aaa aag aga gat gtg gaa ggc att gag gcg ctc atc aca aac Leu Ile Lys Lys Arg Asp Val Glu Gly Ile Glu Ala Leu Ile Thr Asn 180 185 190	576
gcc gag gtc gaa aaa cgg atc ttg gac aga ctt ctg gag aag gga agg Ala Glu Val Glu Lys Arg Ile Leu Asp Arg Leu Leu Glu Lys Gly Arg 195 200 205	624
gcg tat gga aca gac ccg aca cta aag ttc acg cag cac att cag agc Ala Tyr Gly Thr Asp Pro Thr Leu Lys Phe Thr Gln His Ile Gln Ser 210 215 220	672
aag gtg aag ccc gag gtg att gtg aaa atc tac aag gat ttc gtg att Lys Val Lys Pro Glu Val Ile Val Lys Ile Tyr Lys Asp Phe Val Ile 225 230 235 240	720
ccg ctc acg aag aag gtc gaa gtc gac tac ttg ctg aga cgg ctg gag Pro Leu Thr Lys Lys Val Glu Val Asp Tyr Leu Leu Arg Arg Leu Glu 245 250 255	768
gac gag gag gac gat gat gcg acg cag aaa agc ggc ggc tac gtt gac Asp Glu Glu Asp Asp Asp Ala Thr Gln Lys Ser Gly Gly Tyr Val Asp 260 265 270	816
cgg ttt ctc tcc tct ggc ttg tac tag Arg Phe Leu Ser Ser Gly Leu Tyr 275 280	843

<210> 2
 <211> 280
 <212> PRT
 <213> Hansenula polymorpha

 <400> 2

Met Asp Phe Met Lys Pro Glu Thr Val Leu Asp Leu Gly Asn Ile Arg
1 5 10 15

Asp Ala Leu Val Arg Met Glu Asp Thr Ile Ile Phe Asn Phe Ile Glu
20 25 30

Arg Ser Gln Phe Tyr Ala Ser Pro Ser Val Tyr Lys Val Asn Gln Phe
35 40 45

Pro Ile Pro Asn Phe Asp Gly Ser Phe Leu Asp Trp Leu Leu Ser Gln
50 55 60

His Glu Arg Ile His Ser Gln Val Arg Arg Tyr Asp Ala Pro Asp Glu
65 70 75 80

Val Pro Phe Phe Pro Asn Val Leu Glu Lys Thr Phe Leu Pro Lys Ile
85 90 95

Asn Tyr Pro Ser Val Leu Ala Ser Tyr Ala Asp Glu Ile Asn Val Asn
100 105 110

Lys Glu Ile Leu Lys Ile Tyr Thr Ser Glu Ile Val Pro Gly Ile Ala
115 120 125

Ala Gly Ser Gly Glu Gln Glu Asp Asn Leu Gly Ser Cys Ala Met Ala
130 135 140

Asp Ile Glu Cys Leu Gln Ser Leu Ser Arg Arg Ile His Phe Gly Arg
145 150 155 160

Phe Val Ala Glu Ala Lys Phe Ile Ser Glu Gly Asp Lys Ile Val Asp
165 170 175

Leu Ile Lys Lys Arg Asp Val Glu Gly Ile Glu Ala Leu Ile Thr Asn
180 185 190

Ala Glu Val Glu Lys Arg Ile Leu Asp Arg Leu Leu Glu Lys Gly Arg
195 200 205

Ala Tyr Gly Thr Asp Pro Thr Leu Lys Phe Thr Gln His Ile Gln Ser
210 215 220

Lys Val Lys Pro Glu Val Ile Val Lys Ile Tyr Lys Asp Phe Val Ile

gaatagctgc aggcagcgga gagcaggagg acaaccttgg ctctgctgca atggccgaca	780
tcgagtgcct gcagtcgcta tccagaagaa tccattttgg ccgttttgtc gcagaggcta	840
aatttatcag tgagggggac aagattgtgg atctgatcaa aaagagagat gtggaaggca	900
ttgaggcgct catcaciaaac gccgaggctg aaaaacggat cttggacaga cttctggaga	960
aggggaagggc gtatggaaca gacccgacac taaagtccac gcagcacatt cagagcaagg	1020
tgaagcccga ggtgattgtg aaaatctaca aggatttcgt gattccgctc acgaagaagg	1080
tcgaagtcga ctacttgctg agacggctgg aggacgagga ggacgatgat gcgacgcaga	1140
aaagcggcgg ctacgttgac cggtttctct cctctggctt gtactagaaa ttaaaatatt	1200
cagtacttta attattctcg aattctagtt cagataccgc atggtaattt caaaggccag	1260
aaaagtggcc gcgttggtg gggcagctct cagaatagtc ggcgagaatc ctttgactag	1320
ccccaggca ccgctctgtc tccaaatacc cctaatagtc tcaacagcat ttctataaac	1380
cagcttcttg tagttgtccg tctgcatgtt ggacttgatc acatcgatcg gataaatact	1440
gaaccacatc ccgtaacctg ccagcgcccc aaagacgcag agcttccagt tctcgatgtc	1500
cttctgga atattccgag actcgatctc gtttttcacg agagcttcaa aagtcagaaa	1560
atacgctccg ctacccaaac tttctcttgc cagcgtaggt cccagacccc ggtagattaa	1620
cttgatgcct ccggtatggg acagcttctt gatcc	1655

<210> 4
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotide

<400> 4	
aattaaccct cactaaaggg	20

<210> 5
 <211> 22
 <212> DNA
 <213> Artificial

<220>
 <223> Oligonucleotide

<400> 5	
gtaatacgac tcactatagg gc	22

<210> 6
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 6
atatagatct acaaaaacta aacagg

26

<210> 7
<211> 28
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 7
atatagatct gatgcgacgc agaaaagc

28

SEQUENCE LISTING

<110> Gellissen, Gerd
Braus, Gerhard
Pries, Ralph
Krappmann, Sven
Strasser, Alexander

<120> Nucleic Acid Molecule Comprising a Nucleic Acid Coding for a Polypeptide
with Chorismate Mutase Activity

<130> 029474-5007-00

<140> 10/042059

<141> 2001-10-25

<150> DE 199 19 124.7

<151> 1999-04-27

<160> 7

<170> PatentIn version 3.1

<210> 1

<211> 843

<212> DNA

<213> Hansenula polymorpha

<220>

<221> CDS

<222> (1)..(843)

<223>

<400> 1

atg gac ttt atg aag cca gaa aca gtg ctg gac ctt ggc aac att aga	48
Met Asp Phe Met Lys Pro Glu Thr Val Leu Asp Leu Gly Asn Ile Arg	
1 5 10 15	

gat gcc ttg gtc cgg atg gag gat acg atc atc ttc aac ttt atc gag	96
Asp Ala Leu Val Arg Met Glu Asp Thr Ile Ile Phe Asn Phe Ile Glu	
20 25 30	

cgg tcg cag ttc tat gcg tcg ccc tcg gta tac aaa gtc aac cag ttc	144
Arg Ser Gln Phe Tyr Ala Ser Pro Ser Val Tyr Lys Val Asn Gln Phe	
35 40 45	

cct att ccc aac ttc gac ggc tcg ttc ttg gac tgg ctg ttg tcg cag	192
Pro Ile Pro Asn Phe Asp Gly Ser Phe Leu Asp Trp Leu Leu Ser Gln	
50 55 60	

cac gag cga atc cat tcg cag gtg agg aga tac gac gcg cca gac gag	240
His Glu Arg Ile His Ser Gln Val Arg Arg Tyr Asp Ala Pro Asp Glu	
65 70 75 80	

gtg cct ttt ttc ccc aac gtg ctg gaa aaa acg ttt ctg ccc aag atc	288
Val Pro Phe Phe Pro Asn Val Leu Glu Lys Thr Phe Leu Pro Lys Ile	
85 90 95	

aac tac cca tcg gtg cta gcc tcc tac gcg gat gaa atc aac gtc aac	336
Asn Tyr Pro Ser Val Leu Ala Ser Tyr Ala Asp Glu Ile Asn Val Asn	

100

105

110

aaa gag ata ctc aag atc tac acg tca gag ata gta cca gga ata gct 384
Lys Glu Ile Leu Lys Ile Tyr Thr Ser Glu Ile Val Pro Gly Ile Ala
115 120 125

gca ggc agc gga gag cag gag gac aac ctt ggc tcg tgc gca atg gcc 432
Ala Gly Ser Gly Glu Gln Glu Asp Asn Leu Gly Ser Cys Ala Met Ala
130 135 140

gac atc gag tgc ctg cag tcg cta tcc aga aga atc cat ttt ggc cgt 480
Asp Ile Glu Cys Leu Gln Ser Leu Ser Arg Arg Ile His Phe Gly Arg
145 150 155 160

ttt gtc gca gag gct aaa ttt atc agt gag ggg gac aag att gtg gat 528
Phe Val Ala Glu Ala Lys Phe Ile Ser Glu Gly Asp Lys Ile Val Asp

 165 170 175

ctg atc aaa aag aga gat gtg gaa ggc att gag gcg ctc atc aca aac 576
Leu Ile Lys Lys Arg Asp Val Glu Gly Ile Glu Ala Leu Ile Thr Asn
180 185 190

gcc gag gtc gaa aaa cgg atc ttg gac aga ctt ctg gag aag gga agg 624
Ala Glu Val Glu Lys Arg Ile Leu Asp Arg Leu Leu Glu Lys Gly Arg
195 200 205

gcg tat gga aca gac ccg aca cta aag ttc acg cag cac att cag agc 672
Ala Tyr Gly Thr Asp Pro Thr Leu Lys Phe Thr Gln His Ile Gln Ser
210 215 220

aag gtg aag ccc gag gtg att gtg aaa atc tac aag gat ttc gtg att 720
Lys Val Lys Pro Glu Val Ile Val Lys Ile Tyr Lys Asp Phe Val Ile
225 230 235 240

ccg ctc acg aag aag gtc gaa gtc gac tac ttg ctg aga cgg ctg gag 768
Pro Leu Thr Lys Lys Val Glu Val Asp Tyr Leu Leu Arg Arg Leu Glu
245 250 255

gac gag gag gac gat gat gcg acg cag aaa agc ggc ggc tac gtt gac 816
Asp Glu Glu Asp Asp Ala Thr Gln Lys Ser Gly Gly Tyr Val Asp
260 265 270

cgg ttt ctc tcc tct ggc ttg tac tag 843
 Arg Phe Leu Ser Ser Gly Leu Tyr
 275 280

```
<210> 2
<211> 280
<212> PRT
<213> Hansenula polymorpha
```

 $\langle 400 \rangle$ 2

Met Asp Phe Met Lys Pro Glu Thr Val Leu Asp Leu Gly Asn Ile Arg
1 5 10 15

Asp Ala Leu Val Arg Met Glu Asp Thr Ile Ile Phe Asn Phe Ile Glu
20 25 30

Arg Ser Gln Phe Tyr Ala Ser Pro Ser Val Tyr Lys Val Asn Gln Phe
35 40 45

Pro Ile Pro Asn Phe Asp Gly Ser Phe Leu Asp Trp Leu Leu Ser Gln
 50 55 60
 His Glu Arg Ile His Ser Gln Val Arg Arg Tyr Asp Ala Pro Asp Glu
 65 70 75 80
 Val Pro Phe Phe Pro Asn Val Leu Glu Lys Thr Phe Leu Pro Lys Ile
 85 90 95
 Asn Tyr Pro Ser Val Leu Ala Ser Tyr Ala Asp Glu Ile Asn Val Asn
 100 105 110
 Lys Glu Ile Leu Lys Ile Tyr Thr Ser Glu Ile Val Pro Gly Ile Ala
 115 120 125
 Ala Gly Ser Gly Glu Gln Glu Asp Asn Leu Gly Ser Cys Ala Met Ala
 130 135 140
 Asp Ile Glu Cys Leu Gln Ser Leu Ser Arg Arg Ile His Phe Gly Arg
 145 150 155 160
 Phe Val Ala Glu Ala Lys Phe Ile Ser Glu Gly Asp Lys Ile Val Asp
 165 170 175
 Leu Ile Lys Lys Arg Asp Val Glu Gly Ile Glu Ala Leu Ile Thr Asn
 180 185 190
 Ala Glu Val Glu Lys Arg Ile Leu Asp Arg Leu Leu Glu Lys Gly Arg
 195 200 205
 Ala Tyr Gly Thr Asp Pro Thr Leu Lys Phe Thr Gln His Ile Gln Ser
 210 215 220
 Lys Val Lys Pro Glu Val Ile Val Lys Ile Tyr Lys Asp Phe Val Ile
 225 230 235 240
 Pro Leu Thr Lys Lys Val Glu Val Asp Tyr Leu Leu Arg Arg Leu Glu
 245 250 255
 Asp Glu Glu Asp Asp Asp Ala Thr Gln Lys Ser Gly Gly Tyr Val Asp
 260 265 270
 Arg Phe Leu Ser Ser Gly Leu Tyr
 275 280

<210> 3
 <211> 1655

<212> DNA
<213> Hansenula polymorpha

<220>
<221> gene
<222> (1)..(1655)
<223> 1,8 kb genomic DNA-fragment from Hansenula polymorpha

<220>
<221> gene
<222> (1)..(1655)
<223> 1,8 kb genomic DNA-fragment from Hansenula polymorpha

<400> 3
cccggcccaa tgccagcaat atggagacgt ttaggcagaa taggcgttcc atactttctca 60
cgctgcttgt tgccaccgga atatacaccg cattgcagtt tgcacacatc atactatatg 120
acgattacat tggcgggaacg tatcgcgagt cgctcacgag acgcattaga atgacagaga 180
aatcgcgaaa cgaccttata gacgcacgtg aaaactacgg gtttggaggc agcaaggagg 240
agcgaatcca gcggtttttg tggttcagac atctttcgtg gcttttaggc gaggataagc 300
gaacttgagg agcggtttttt ttttcctggt tagtttttgt aggtatggac tttatgaagc 360
cagaaacagt gctggacctt ggcaacatta gagatgcctt ggtccggatg gaggatacga 420
tcactttcaa ctttatcgag cggtcgcagt tctatgcgtc gccctcggta taaaaagtca 480
accagttccc tattcccaac ttcgacggct cgttcttgga ctggctgttg tcgcagcacg 540
agcgaatcca ttcgcaggtg aggagatacg acgcgccaga cgaggtgcct tttttcccca 600
acgtgctgga aaaaacgttt ctgcccaaga tcaactaccc atcgggtgcta gcctcctacg 660
cggatgaaat caacgtcaac aaagagatac tcaagatcta cacgtcagag atagtaccag 720
gaatagctgc aggcagcgga gagcaggagg acaaccttgg ctctgctgca atggccgaca 780
tcgagtgcct gcagtcgcta tccagaagaa tccattttgg ccgttttgtc gcagaggcta 840
aatttatcag tgaggggggac aagattgtgg atctgatcaa aaagagagat gtggaaggca 900
ttgaggcgct catcaciaac gccgaggtcg aaaaacggat cttggacaga cttctggaga 960
aggggaagggc gtatggaaca gacccgacac taaagttcac gcagcacatt cagagcaagg 1020
tgaagcccga ggtgattgtg aaaatctaca aggatttcgt gattccgctc acgaagaagg 1080
tcgaagtcga ctacttgctg agacggctgg aggacgagga ggacgatgat gcgacgcaga 1140
aaagcggcgg ctacgttgac cggtttctct cctctggctt gtactagaaa ttaaaatttt 1200
cagtacttta attattctcg aattctagtt cagataccgc atggtaattt caaaggccag 1260
aaaagtggcc gcgttggtg gggcagctct cagaatagtc ggcgagaatc ctttgactag 1320
ccccaggca ccgctctgtc tccaaatacc cctaatagtc tcaacagcat ttctataaac 1380
cagcttcttg tagttgtccg tctgcatggt ggacttgatc acatcgatcg gataaatact 1440

gaaccacatc ccgtaacctg ccagcgcccc aaagacgcag agcttccagt tctcgatgtc 1500
cttcctggca atattccgcg actcgatctc gtttttcacg agagcttcaa aagtcagaaa 1560
atacgctccg ctacccaaac tttctcttgc cagcgtaggt cccagacccc ggtagattaa 1620
cttgatgcct cccgtatggg acagcttctt gatcc 1655

<210> 4
<211> 20
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 4
aattaaccct cactaaaggg 20

<210> 5
<211> 22
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 5
gtaatacgac tcactatagg gc 22

<210> 6
<211> 26
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 6
atatagatct acaaaaacta aacagg 26

<210> 7
<211> 28
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide

<400> 7
atatagatct gatgcgacgc agaaaagc 28